

本周周报（2013.10.21-2013.10.27）

郭方舟

本周工作

1. 空气污染数据可视化

1) 在上周计算的信息熵的基础上，对计算进行了改进。

上周分别计算了以 1 天，3 天，1 周，2 周，1 月为窗长的信息熵，但是这种计算熵的方式没有办法做出堆叠三角形的效果。本周将计算方式进行修改，我们将窗长选择调整为从五个小时开始逐一增加，直到窗长等于 365 天；而窗移则设定为一个小时。由于计算到 365 天时这个三角形非常巨大，因此我们暂时以 96 小时为最大窗长。

2) 将层级计算的信息熵用堆叠三角形的形式可视化。

得到信息熵之后，以堆叠三角形的方式对信息熵进行了可视化，效果如下图所示：



这里，使用灰度代表每一个点计算出的熵值。从上到下代表窗长依次增加。从截图中可以看到，这种可视化方法可以到位的表达出信息熵随时间增加的变化方式以及其中明显的 pattern。比如：

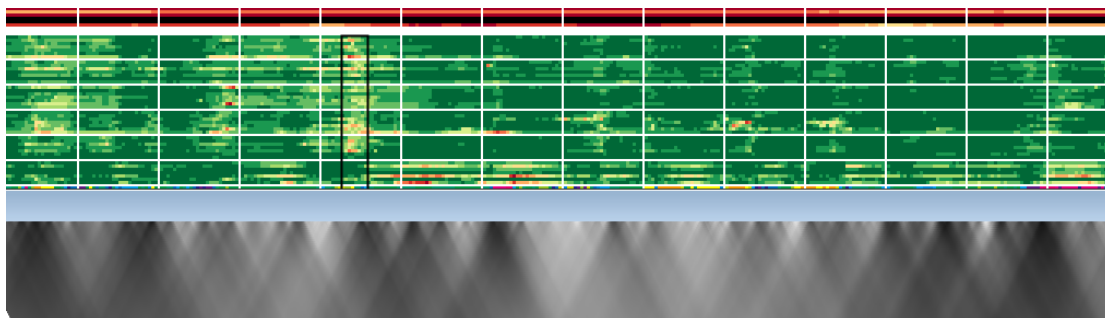


这里有一个明显的浅色三角，这代表这一段时间里不管是从小尺度还是从大尺度来看，信息熵都非常小，也就是说这段时间的天气变化不是很大。



这张截图虽然整体灰度比较浅，但是可以看到在窗长较小的时候，其灰度值是比较大的，这代表着有一些小尺度的扰动出现。

3) 重构代码，将像素图和堆叠图连通，并做了基本的交互。



由于像素图和堆叠图的面积都相当大，都需要通过交互才能看到整张图，所以将两个视图放在一个窗口下是不太现实的，考虑到将来可能有更多的功能要添加到系统中，所以将像素图作为主窗口，而堆叠图作为一个对话框进行整合。在堆叠图的一个点进行点击后，会在像素图上以一个黑色矩形标出这个点所代表的数据点。

从目前的效果来看，使用信息熵对该数据进行处理能够得到比较好的效果。上一次讨论的三个工作中的第一个工作，自适应表达的基础基本完成。下一步工作是将像素图中的不包含事件的时间段合并，结合信息熵和堆叠图，可以有两种实现方向：1. 计算出熵值相对低的时间段并将这些时间段合并；2. 在堆叠图上进行交互，让用户在图上划线，根据用户的选择将某些时间段合并起来。将无时间的时间段合并之后，事件相对的也就凸显出来了。

目前的实现方式仍具有一定的问题：将所有属性合并到一起计算，这使得单属性的横向事件可能无法在堆叠图上体现。这个问题可以通过将属性分割，单属性计算信息熵，但是这种方式会导致堆叠图的大小进一步增大，增加分析难度。

下周工作

1. 空气污染数据的可视化

下一周应该要将自适应表达这工作完成，并将相关的交互完成，同时开始进行第二项任务 **color map** 的交互调整。

2. 完成秋学期课程作业

秋学期即将结束，需要完成马克思与当代这门课的课程论文以及计算机图形学的第一个作业。